La fonction d'entropie

L'objet de ce problème est l'étude de quelques propriétés élémentaires de la fonction d'entropie qui mesure l'incertitude sur la valeur prise par une variable aléatoire donnée.

NOTATIONS : Dans tout le problème n désigne un entier naturel.
On note $I_n = \{k \in \mathbb{N};\ 0 \leq k \leq n\}$ l'ensemble des entiers naturels compris entre 0 et n.
Si f et g sont deux fonctions, on note, lorsque cela a un sens, $f \circ g$ leur composée.

1. **Préliminaire.**
   (a) On note h l'application de $\mathbb{R}^+$ dans $\mathbb{R}$ définie par $h(0) = 0$ et pour tout $x > 0$, $h(x) = x\ln(x)$.
   Montrer que h est continue sur $\mathbb{R}^+$ et tracer l'alure de son graphe en précisant la tangente au point d'abscisse 1.
   (b) Montrer que, pour tout $x > 0$, $\ln(x) \leq x - 1$ et que $\ln(x) = x - 1$ si et seulement si $x = 1$.

Si, n étant un entier, X est une variable aléatoire discrète à valeurs dans I_n, de loi $p_i = \mathbb{P}(X = i)$ pour tout entier $i \in I_n$, on note

$$H(X) = -\sum_{i=0}^{n} h(p_i).$$

2. **Exemple.**
   Dans cette question, X est une variable aléatoire suivant une loi uniforme sur I_n. Calculer $H(X)$.

3. **Propriétés de l'entropie**
   Dans cette question X_0 est une variable aléatoire de loi uniforme sur I_n et X désigne une variable aléatoire discrète (=définie sur un univers porbabilisé fini) à valeurs dans I_n.
   (a) Montrer que $H(X) \geq 0$ puis que $H(X) = 0$ si et seulement si X est une variable aléatoire certaine i.e. : $\exists! i_0 \in I_n;\ p_{i_0} = 1$ et $\forall i \neq i_0,\ p_i = 0$.
   (b) En utilisant le préliminaire, montrer que pour tout $i \in I_n$, $-h(p_i) + p_i \ln\left(\frac{1}{n+1}\right) \leq \frac{1}{n+1} - p_i$.
   (c) En déduire que $H(X) \leq H(X_0)$, puis que $H(X) = H(X_0)$ si et seulement si X suit la loi uniforme (par l'absurde on montrera que si X ne suit pas la loi uniforme alors $H(X) < H(X_0)$).
   (d) Donner une interprétation de ce résultat.

4. **Entropie d'un couple.**
   Dans cette question n et m sont des entiers naturels non nuls, (X, Y) et (X', Y') sont deux couples de variables aléatoires discrètes.
   Les variables aléatoires X et X' (resp. Y et Y') sont à valeurs dans I_n (resp. I_m).
   Pour tout $i \in I_n$, on note p_i la probabilité $p_i = \mathbb{P}(X = i)$, pour tout $j \in I_m$, on note q_j la probabilité $q_j = \mathbb{P}(Y = j)$.
   Enfin, pour tout couple $(i, j) \in I_n \times I_m$, on note λ_{ij} la probabilité $\lambda_{ij} = \mathbb{P}(X = i; Y = j)$ et μ_{ij} la probabilité $\mu_{ij} = \mathbb{P}(X' = i; Y' = j)$.
   On suppose que, pour tout (i, j) de $I_n \times I_m$, $\lambda_{ij} \neq 0$ et $\mu_{ij} \neq 0$.

NOTATIONS : On définit l'entropie du couple (X, Y) par

$$H(X, Y) = -\sum_{i=0}^{n}\sum_{j=0}^{m} h(\lambda_{ij}).$$

et l'information entre les couples (X, Y) et (X', Y') par

$$K(X, Y, X', Y') = -\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda_{ij} \ln\left(\frac{\mu_{ij}}{\lambda_{ij}}\right).$$

(a) **Propriétés de l'information entre deux couples**.

i. Rappeler les valeurs de $\sum_{i=0}^{n}\sum_{j=0}^{m}\lambda_{ij}$ et de $\sum_{i=0}^{n}\sum_{j=0}^{m}\mu_{ij}$. En déduire que : $K(X,Y,X',Y') = -\sum_{i=0}^{n}\sum_{j=0}^{m}\lambda_{ij}\left[\ln\left(\frac{\mu_{ij}}{\lambda_{ij}}\right)-\frac{\mu_{ij}}{\lambda_{ij}}+1\right]$.

ii. A l'aide de la question 1., établir que $K(X,Y,X',Y')\geq 0$ et qu'on a égalité si et seulement si les deux couples (X,Y) et (X',Y') ont même loi conjointe.

iii. On suppose que les deux variables aléatoires X' et Y' sont indépendantes, respectivement de même loi que X et Y. Pour tout i de I_n, on notera $p_i=\mathbb{P}(X=i)=\mathbb{P}(X'=i)$, respectivement pour tout j de I_m, $q_j=\mathbb{P}(Y=j)=\mathbb{P}(Y'=j)$.
Montrer que $K(X,Y,X',Y')=H(X)+H(Y)-H(X,Y)$.
Déduire de ce qui précède que $H(X,Y)\leq H(X)+H(Y)$ (inégalité (∇)).
Donner une condition nécessaire et suffisante pour que cette inégalité soit une égalité.

REMARQUE : *l'inégalité* (∇) *a été obtenue en supposant, pour tout* (i,j) *de* $I_n\times I_m$, $\lambda_{ij}\neq 0$ *et* $\mu_{ij}\neq 0$. *On admet qu'elle demeure vraie sans cette condition.*

(b) **Entropie conditionnelle.**

On définit l'entropie conditionnelle de Y sachant X par $H(Y|X)=H(X,Y)-H(X)$.

Elle mesure l'incertitude restant sur la valeur de Y*, la valeur de* X *étant connue.*

i. Montrer que $H(Y|X)\leq H(Y)$.

ii. On considère $m+1$ réels a_0, a_1, ..., a_m compris entre 0 et 1.

A. Dans cette question, on suppose $(a_0,a_1,\dots,a_m)\in]0,1]^{m+1}$.
Montrer que pour tout i de I_m, $\ln(a_i)\leq\ln(a_0+a_1+\cdots+a_m)$.
En déduire que $\sum_{i=0}^{m}h(a_i)\leq h\left(\sum_{i=0}^{m}a_i\right)$. (inégalité $(*)$)

B. L'inégalité $(*)$ demeure-t-elle si $(a_0,a_1,\dots,a_m)\in[0,1]^{m+1}$? (on posera $\widetilde{I}_m=\{i\in I_m;\ a_i\neq 0\}$)

C. Montrer que l'inégalité $(*)$ est une égalité si et seulement s'il existe au plus un indice $i\in I_m$ pour lequel $a_i\neq 0$.

(c) Montrer que pour tout $i\in I_n$, $\sum_{j=0}^{m}h(\lambda_{ij})\leq h(p_i)$. En déduire que $H(Y|X)\geq 0$.

5. **Une application.**

Un jeu oppose deux joueurs. Le joueur A tire une boule au hasard dans une urne contenant 2009 boules numérotées de 0 à 2008. On note Y la variable aléatoire égale au numéro de la boule tirée. Dans le but de déterminer Y, le joueur B pose une série de N question(s) où N est un entier naturel non nul fixé à l'avance. Ces questions sont impérativement de la forme : "le nombre choisi appartient-il à E ?" où E est un sous-ensemble de I_{2008} pouvant varier à chaque question.
Le joueur A répond par l'affirmative si $Y\in E$, par la négative sinon.
Au bout des N question(s), B donne sa réponse.
En cours de jeu, on note $X_i=1$ si la réponse de A est affirmative à la $i$ème question de B et $X_i=0$ sinon.

(a) On pose $X=\sum_{i=1}^{N}X_i 2^{i-1}$. On admet que X est une variable aléatoire. Montrer qu'elle est à valeurs dans $I_{(2^N-1)}$ (on ne cherchera pas à déterminer sa loi).

(b) Déterminer l'entropie $H(Y)$.

(c) Montrer que $H(X)\leq N\ln(2)$ (on utilisera l'inégalité vue à la question 3.(c)).

(d) Expliquer, en langage courant, pourquoi on a $H(X|Y)=0$. En déduire que $H(X,Y)=H(Y)$.

(e) Montrer que $H(Y|X)\geq\ln(2009)-N\ln(2)$.

(f) Le joueur B prétend pouvoir trouver la réponse à coup sûr en 10 questions (ou moins). Qu'en pensez-vous ? En combien de questions pouvez-vous donner à coup sûr la valeur de Y ?
(On utilisera l'inégalité démontrée à la question précédente.)